STOCHASTIC STREAMFLOW ANALYSIS AND SIMULATION USING DISAGGREGATION MODEL

AZWAN BIN MUSTAPHA

A project submitted in partial fulfillment of the
requirement for the award of the degree of
Master of Engineering (Civil - Hydraulic and Hydrology)

Faculty of Civil Engineering
University Teknologi Malaysia

JANUARY 2012

To my beloved parent (Ma and Bat)

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious and the Most Merciful

Alhamdulillah, all praises to Allah for the strengths and His blessing that made the author possible to complete this thesis in the time given.

Foremost, I would like to express my sincere gratitude to my supervisor Assoc. Prof. Dr. Sobri Bin Harun for the continuous support of my Mater’s Project, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

Special appreciation goes to the thesis panel member, for providing comments and suggestion in completing this study. Last but not least, I would like to thank to my beloved father and family for their support and encouragement for me to finish my Master degree.

ABSTRACT

Synthetic hydrological series is beneficial for the water resources planning, management and design purposes such as evaluating water supply management decision and reservoir design. Stochastically generated data can be used in any fields of study that use historical data. It would produce a better result for any application where the result depends on the time series pattern. Therefore, this study examines three stochastic disaggregation models that are capable of reproducing statistical parameters especially means and standard deviation of historical data series. The objectives of this study are to test, identify and confirm the best stochastic disaggregation model in generating synthetic data series. Stochastic Analysis Modeling and Simulation (SAMS-2000) is used to generate the synthetic hydrological data series. The method is applied for single site cases and comparison is made between three disaggregation models namely Valencia and Schaake model, Mejia and Rouselle model, and Lane model. The simulations of monthly streamflow are carried out for five stations from Kedah, Perak and Selangor. The comparison of results shows that Valencia and Schaake (VLSH) model is the most satisfactory and robust model that preserves both monthly and annual statistical parameters of the historical data sequences. This is true for both untransformed and transformed series. Therefore, it is recommended to use VLSH model for simulation of water resources in study area.

ABSTRAK

Siri hidrologi sintetik adalah sangat bermanfaat dalam merancang pengurusan sumber air dan tujuan reka bentuk seperti menilai keputusan dan pengurusan bekalan air mentah serta reka bentuk sistem takungan. Data yang dihasilkan secara stokastik boleh digunakan di dalam semua bidang yang menggunakan data asal. Ia akan menghasilkan keputusan yang lebih baik memandangkan keputusan bergantung kepada corak siri masa,. Oleh yang demikian, kajian ini akan menilai tiga model disagregasi yang berkebolehan dalam mengekalkan ciri - ciri statistik data asal terutamanya purata aliran sungai dan sisihan piawai. Objektif utama kajian ini adalah untuk mengenal pasti model disagregasi yang terbaik dalam menjana data sintetik. Perisian "Stochastic Analysis Modeling and Simulation" (SAMS-2000) akan digunakan untuk menjana data sintetik. Tiga disagregasi model iaitu Valencia dan Schaake (VLSH), Mejia dan Rouselle (MJRS) serta Lane (LANE) digunakan dan perbandingan dibuat bagi kes tapak kajian tunggal. Simulasi telah dilakukan bagi lima stesen kajian di Kedah, Perak dan Selangor. Hasil kajian menunjukkan bahawa model VLSH adalah model yang paling terbaik dalam mengekalkan ciri - ciri statistik data bulanan dan tahunan siri aliran sungai. Ini adalah benar bagi siri tanpa transformasi dan siri transformasi. Oleh yang demikian, model VLSH adalah disarankan dalam menjalankan simulasi sumber air ke atas kawasan kajian.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

A, B and C	Parameter matrices
Q	Singel value and a column vector which contains an annual value
Q_t	Matrix of the annual streamflow value of year t
Q_v	Annual series vector
$Q_{v,t}$	Generated annual streamflow
m	Rank of residual matrix
M_o	Lag - zero correlation matrix population moment
M^1	Lag - one correlation matrix or population moment
M^{-1}	Inverse matrix of population
MT	Transpose matrix of population moment
VLSH	Valencia and Schaake model
MJRS	Mejia and Rouselle model
LANE	Lane model

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

1.1 Background

Stochastic simulation of water resources time series in general and hydrologic time series in particular has been widely used for several decades for various problems related to planning and management of water resources systems. Typical examples are determining the capacity of a reservoir, evaluating the reliability of a reservoir of a given capacity, evaluation of the adequacy of a water resources management strategy under various potential hydrologic scenarios, and evaluating the performance of an irrigation system under uncertain irrigation water deliveries (Salas et al, 1980; Loucks et al, 1981)

Stochastic simulation of hydrologic time series such as streamflow is typically based on mathematical models. For this purposes a number of stochastic models have been suggested in literatures (Salas, 1993; Hipel and McLeod, 1994). Using one type of model or another for a particular case at hand depends on several factors such as, physical and statistical characteristics of the process under consideration, data availability, the complexity of the system, and the overall purpose of the simulation study (Salas et al.1960). Instead of well defined mathematical models, appropriate modeling strategies and schemes also needed for complex water resources system (Salas et al. 2006). Given the historical record, one would like the model to reproduce the historical statistics. Once a model has been selected, the next step is to estimate the model parameter, then to test whether the model represents reasonably well the process under consideration, and finally to carry out the needed simulation study.

Hydrological data such as flows and rainfall are the basic information used for the design of water resources systems. Hence, the studies are needed regarding the accuracy of the data required for various types of water resources planning and management. This situation needs one to select the most appropriate time interval of hydrological data for the design purposes because the characteristics behavior between time intervals is different to each other. Despite the resolution of the time interval, the data quality and accuracy should also be taken into account. The data uncertainties and randomness that is one of the factors that stems from difficulties in estimating future demand for water development.

Many hydrologist use the forecasting technique to design and manage the water resources system. The purpose of forecasting is to use the time series model fitted to a data set to obtain the most accurate estimation or prediction of the future unknown series. However, forecasting itself is unable to test whether or not a class of time series model statistically conserves the historical characteristics of the data sets to which the set of model is fitted (Box and Jenkins, 1976).

Simulation methods for the hydrologic time series can be classified into disaggregation and aggregation methods. Disaggregation models are generally considered are very reliable feature for flow and rainfall simulation. The first wellaccepted model was presented by Valencia and Schaake (1973) by developing Valencia - Schaake (VLSH) model. Further studies, modification and applications of disaggregation model introduce the extended model developed by Mejia and Rouselle (1976) which is Mejia - Rousselle (MJRS) model and Lane (1980) which is LANE model. Aggregation model is a new technique by which an assumption is made regarding the basic form of the original or lower level time series are calculated. In order to model adequately the seasonally varying correlation structure and to preserve the stationary statistical properties within each season, one would have to consider the families of the periodic autoregressive (PAR) model or periodic autoregressive and moving average (PARMA) model. The application of these models has been attractive in simulation area mainly because, the form has an intuitive type of theme dependence and they are simplest model to use.

Simulation of hydrological time series such as streamflow and rainfall has been attractive because of several reasons. One is the particular nature of hydrological process in which periodic properties are important in the mean, variance, covariance and skewness. Another one is that some hydrologic time series include complex characteristic such as long term dependence and memory. Still another one is many of the stochastic models useful in hydrology and water resources have been developed specifically oriented to fit the needs of water resources, for instance temporal and spatial disaggregation models.

This study will only focus on streamflow simulation based on the disaggregation procedure using Valencia - Schaake (VLSH) model, Mejia - Rouselle (MJRS) model and Lane (LANE) model. The stochastic simulation models develop by Salas et.al (1996) was used to generate the synthetic series from historical records. The effectiveness of the models depends on the estimation of model parameter, fitting stages and diagnostic check. The model estimate stage needs to be check in order to verify how well it represents the historical flow and rainfall series. The evaluation of the selected model is based on the preservation of the statistical characteristics such as mean, standard deviation and skewness. Therefore, it is necessary to evaluate the validity of a model before it is used for such purpose.

1.2 Statement of Problems

Water is the main source and very important for the living being to continue and support their life. However, problem related to water become crucial nowadays such as water shortage, drought, floods and many others. This happened due to ineffective water resources management, design and planning. The proper management is important in order to sustain the water resource available. A good planning and design of water resources system required a good and reliable data series. Instead of using the historical data that contain randomness and errors, a stochastic simulation approach is applied to obtain a much reliable and good synthetic data.

Stochastic data generation aims to provide alternative hydrological data sequences that are likely to occur in the future. These data sequences (rainfall and streamflow) particularly monthly data series are widely used in water resources for short and long term planning and management. In reality, these data contain randomness and uncertainty. The effects may caused by human or nature changed. A stochastic model must be adopted as this problem cannot be solved analytically.

Various stochastic processes have been developed by the researcher in the literature such as aggregation and disaggregation. These models used the historical data to estimate the model parameter and generating new series from the estimated parameter.

Disaggregation models are currently being used widely to generate synthetic data that preserve the historical statistic parameters (mean, standard deviation, skewness, .etc). The generated series are evaluated from the preservation of the historical properties. Disaggregation models produce monthly data sequence by disaggregating the annual data that have been generated by suitable annual data simulation.

The stochastic simulation is a potential method in solving the natural process of streamflow series based on the statistical characteristics. This technique can model the random component in the system. This study will help in identify the right model to represent the time series that preserved the historical properties.

1.3 Objectives

The objectives of this study are:

i. To test various stochastic disaggregation models capable of reproducing the historical statistical parameters especially mean, standard deviation and skewness of the streamflow series.

ii. To identify and confirm the best disaggregation model in streamflow simulation.

1.4 Scope of Study

The study covers the application of single site disaggregation methods for monthly streamflow simulation. The disaggregation models used in this study are Valencia and Schaake (VLSH), Mejia and Rouselle (MJRS) and Lane (LANE) model. Stochastic Analysis Simulation and Modeling (SAMS-2000) package is used to simulate the monthly streamflow in this study.

The data analysis of the generated synthetic series comprises of time series plot and Box - Whisker plots. The results of generated series will be examined and compared with historical mean, standard deviation and skewness coefficient to identify the best model. The best model is identified when it preserves the historical statistical parameters satisfactorily. The streamflow gauging stations are selected from three state namely; Kedah, Perak and Selangor. Table 1.1 shows the details of the hydrological data site selected.

Table 1.1: The details of the hydrological data sites

Station Id	State	River Basin	Station Name	Record Length
5806414	Kedah	Muda	Sg. Muda at Jeniang	1960 - 2009 (50 years)
5106433	Perak	Sg. Kerian	Sg. Ijok at Titi Ijok	1960 - 2008 (49 years)
5206432	Perak	Sg. Kerain	Sg. Kerian at Selama	1960 - 2009 (50 years)
3813411	Selangor	Bernam	Sg. Bernam at SKC	1961 - 2009 (49 years)
3615412	Selangor	Bernam	Sg. Bernam at Tanjung Malim	1960 - 2009 (50 years)

1.5 The Important of Study

This study will present the disaggregation models for generation of alternative sequence of monthly and annual hydrologic data sequences. The proposed model is therefore a valuable tool for flow simulation studies around Malaysia. The synthetic data is essential parameter for development of proper water resources planning and management. The results of this study also provide a new approach for solving t` problem in the historical data series.

REFERENCES

Bojilova, E. K., (2004) Disaggregation Modeling of spring Discharges, *Int.J.Speleol*, 33(1/4): 65 - 72

Box, G.E.P., and Jenkins, G.M. (1976). *Time Series Analysis Forecasting and Control*, Holden-Day Inc, California

Box, G.E.P., Jenkins, G.M., and Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control*. Third edition, Prentice Hall.

Frevert, D.K., (2002) Mathematical Models of Small Watershed Hydrology and Application, *Water Resources Publication*, LLC

Frick, D. M. and Salas, J. D. (1991*). Evaluating Modeling Strategies for a Complex Water Resources System. Hydrology for the Water Management of Large River Basin,* IASH, Publ. No. 201.

Gabric, I. and Kresic, N. (2009) Time Series Models. N. Kresic (Ed). *Groundwater Resources: Susutainability, Management and Restoration*, McGraw Hill, New York (2009). Pp. 657-661.

Hipel, K.W., McBean, E. A., and McLeod, A. I. (1979). Hydrological Generating Model Selection. *Journal of Water Resour. Plang. And Mgmt. Division*. 105:223-242

Hipel, K.W., and McLeod, A.I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam.

Kumar, D. N., Lall, U., and Petersen, M.R. (2000). Multisite Disaggregation of Monthly to Daily Streamflow. *Water Resources Research* 36(7):1823-1833

Lane, W.L. (1979). *Applied Stochastic Technique (LAST computer package), User Manual*. Division of Planning Technical Services, Bureau of Reclamation, Denver, Colorado USA.

Lane, W. L. (1982). *Corrected Parameter Estimates for Disaggregation Schemes in Statistical Analysis of Rainfall and Runoff*. Water Resources Publication, Littleton

Langouisis, A. and Koutsoyiannis, D. (2006). A Stochastic Methodology for Generation of Seasonal Time Series Reproducing Overyear Scaling Behavior, *Journal of Hydrology* 322(2006) 138 - 154

Lee, T., Salas, J.D, and Praire, J. (2010). An enhanced non-parametric streamflow disaggregation models with genetic algorithm. *Water Resource Research*

Loucks, D.P., Stedinger, J.R., and Haith, D.A. (1981). *Water Resource Systems Planning and Analysis*. Prentice-Hall, Englewood Cliffs, New York.

Maheepala, S., and Perera, B.J.C. (1996). Monthly Hidrologic Data Generation by Disaggregation. *Journal of Hydrology*. 178:277-291

McCuen, R. H. amd Snyder, W. M. (1986). Hydrologic Modeling, Statistical Methods and Applications. Prentice - Hall. Englewood Cliffs. NJ.

McLeod, A.I. (1995). *Diagnostic Checking Periodic Autoregression Models With Application*. Journal of Time Series Analysis. 15(2), 221-233.

Mejia, J.M., and rouselle, J. (1976). Disaggregation Models in Hydrology Revisited. *Water Resour. Res.* 12(2):185-186

Mondal, M. S. and Wasimi, S. A. (2006). Generating and Forecasting Monthly Flows of Ganges River with PAR model, *Journal of Hydrology*. 323(1-4):41-56

Salas, J. D., Delleur, J. W., Yevjevich, V., and Lane, W.L. (1980). *Applied Modeling of Hydrological Time Series.* Water Resources Publication. Littleton, Colorado.

Salas, J.D. (1993). *Analysis and Modeling of Hydrologic Time Series*. In: Maidment,D., (Ed.).Handbook of Hydrology. McGraw Hill, New York

Salas, J.D. and Smith, R. A. (1980). *Uncertainties in Hydrologic Time Series Analysis.* Preprint No. 80 - 158, ASCE Convention and Exposition Portland, Oregon

Salas, J.D., Marco, J.B., Harboe, R., (1993) *Stochastic Hydrology and Its Use in Water resources System Simulation and Optimization*, Kluwer Academic Publication, Netherlands

Salas, J.D., Saada, N. M., Lane, W. L., and Frevert, D. K. (1996). *Stochastic Analysis Modeling, and Simulation (SAMS) Version 96.1 User's Manual.* Technical Report No.8, Hydrological Science and Eng. Program, Eng. Res. Center, Col State.

Salas, J.D., Sveinsson, O.G., Lane, W.L., Frevert, D.K. (2006). *Stochastic streamflow simulation using SAMS - 2003*. Journal of Irrigation and Drainage Engineering

Salas, J.D., Sviensson, O.G.B., Lane, W. L., and Frevert, D.K. (2007). *Stochastic Analysis, Modeling and Simulation (SAMS)* Version 2007-User's Manual.

Thomas, H.A., and Fiering, M.B. (1962). *Mathematical Synthesis of Streamflow Sequences for the Analysis of River Basins by Simulation*. In: Maass, A., Hufschimidt,M.M., Dorfman, R., Thomas, H. A., Marglin, S.A., and Fair,G. M. (eds.) Design of Water Resource. Cambridge, Massachusetts, Harvard University

Valencia, R.D., and Schaake. J.C. (1973). *Disaggregation processes in stochastic hydrology*. Water Resources Research, 9(3)

Wang, W (2006). *Stochastically, Nonlinearity and Forecasting of Streamflow Process*, ISO press

Yevjevich, V. (1963). *Fluctuation of Wet and Dry Years Part 1: Research Data Assembly and Mathematical Models*. Hydrology Paper, Colorado State University , FortCollins, Colo.